N86 - 28281

Fast Frequency Acquisition via Adaptive Least Squares Algorithm

R. Kumar[1]

Communications Systems Research Section

A new least squares algorithm is proposed and investigated for fast frequency and phase acquisition of sinusoids in the presence of noise. This algorithm is a special case of more general, adaptive parameter-estimation techniques. The advantages of the algorithms are their conceptual simplicity, flexibility and applicability to general situations. For example, the frequency to be acquired can be time varying, and the noise can be non-gaussian, non-stationary and colored.

As the proposed algorithm can be made recursive in the number of observations, it is not necessary to have a priori knowledge of the received signal-to-noise ratio or to specify the measurement time. This would be required for batch processing techniques, such as the fast Fourier transform (FFT). The proposed algorithm improves the frequency estimate on a recursive basis as more and more observations are obtained. When the algorithm is applied in real time, it has the extra advantage that the observations need not be stored. The algorithm also yields a real time confidence measure as to the accuracy of the estimator.

I. Introduction

The problem of estimating the parameters of a sinusoidal signal has received considerable attention in the literature, see for example Refs. 1–7 and their references. Such a problem arises in diverse engineering situations such as carrier tracking for communications systems and the measurement of Doppler in position location, navigation and radar systems.

A variety of techniques have been proposed in the literature to solve such problems including, to mention a few, the appli-

cation of the fast Fourier transform (FFT) (as in Refs. 1, 2), one and two dimensional Kalman filters based on a linearized model (Ref. 5), a modified extended Kalman filter that results in a phase locked loop (Ref. 6), or a digital phase locked loop derived on the basis of linear stochastic optimization (Ref. 7).

The fact that there are so many different techniques to solve the problem indicates the importance of the problem. This, however, also implies that there is no single technique superior to all others in all possible situations and/or with respect to different criteria such as computational complexity, statistical efficiency, etc.

---

[1] Also with the faculty of the Electrical Engineering Department at California State University, Long Beach.

In this article we propose the application of the least squares parameter estimation technique to the estimation of an unknown frequency. The least squares algorithm has been extensively studied in the literature in terms of convergence, computational requirements, etc. (Refs. 8, 9), and has found varied applications in a wide variety of communication and signal processing problems. This is due to the relative simplicity of the least squares algorithm and its attractive convergence rates. In Ref. 9 for example, it has been shown that the algorithm exhibits an initial factorial convergence rate followed by exponential convergence. Such a convergence is very desirable in almost all estimation situations including the one under consideration.

When the least squares (LS) algorithm is implemented via the fast algorithm of Ref. 10, the computational requirements of the algorithm compare favorably to the FFT algorithm. The least squares algorithm offers, in addition to the above discussed rapid initial convergence, several other desirable features. First the least squares algorithm provides final estimates of frequency, whereas FFT estimation requires use of a secondary algorithm to interpolate between frequencies. Secondly, using an exponentially weighted least squares algorithm, it is possible to track a time varying frequency. We compare the least squares algorithm to the FFT since the latter is "close" to the optimum in terms of the statistical efficiency (Ref. 1).

In Section II we present the signal model followed by the least squares algorithm in Section III. Section IV analyzes the estimation error of the algorithm. In Section V a few simulation examples are presented. The last section of the article contains some concluding remarks.

II. The Signal Model

Consider the problem of estimating an unknown frequency w_d from the measurements y_k, z_k below

$$y_k = A \sin(w_d t_k + \phi) + n_{ik}$$
$$z_k = A \cos(w_d t_k + \phi) + n_{qk}, \qquad k = 1, 2, \ldots \tag{1}$$

Here the sequence $\{y_k, z_k\}$ represents the samples of the in-phase and quadrature components of a received signal $s(t)$ obtained by demodulating $s(t)$ by a carrier reference signal $r(t)$ and its 90-deg phase shifted version respectively, i.e.,

$$s(t) = A \sin(w_0 t + \phi_0) + n(t)$$
$$r(t) = 2 \sin(w_c t + \phi_c), \qquad \phi = \phi_0 - \phi_c, w_d = w_c - w_0$$

with n_{ik} and n_{qk} denoting the samples of the quadrature components of white noise $n(t)$. The algorithm can be easily extended to the case where $n(t)$ is a colored noise.

With a power series expansion for the sine and cosine functions, the measurement equations can be written in alternative forms as follows:

$$y_k = A \sin(w_d t_k) \cos\phi + A \cos(w_d t_k) \sin\phi + n_{ik}$$
$$z_k = A \cos(w_d t_k) \cos\phi - A \sin(w_d t_k) \sin\phi + n_{qk}$$

or

$$
\begin{bmatrix} y_k \\ z_k \end{bmatrix} =
\begin{bmatrix}
A \sin\phi & A \cos\phi\, w_d & -\dfrac{A \sin\phi}{2!} w_d^2 & -\dfrac{A \cos\phi}{3!} w_d^3 & \cdots & +\dfrac{A \sin\phi}{(n-1)!} w_d^{n-1} \\
A \cos\phi & -A \sin\phi\, w_d & -\dfrac{A \cos\phi}{2!} w_d^2 & +\dfrac{A \sin\phi}{3!} w_d^3 & \cdots & +\dfrac{A \cos\phi}{(n-1)!} w_d^{n-1}
\end{bmatrix}
\begin{bmatrix} 1 \\ t_k \\ t_k^2 \\ \cdot \\ \cdot \\ \cdot \\ t_k^{(n-1)} \end{bmatrix}
+ \begin{bmatrix} n_{ik} \\ n_{qk} \end{bmatrix} \tag{2}
$$

In the above approximation the terms of the order $(w_d t_k)^n / n!$ and smaller order have been ignored (assuming here that $w_d t_k < n$). With obvious definitions, the measurement equation can be written in a form "linear in parameters."

$$Z_k = \theta' x_k + n_k \tag{3}$$

In the above, the prime $(')$ denotes transpose, $Z_k' = [y_k\, z_k]$, $n_k' = [n_{ik}\, n_{qk}]$, x_k' denotes the observable state vector $[1\ t_k$

$t_k^2 \ldots t_k^{n-1}$] and θ' is the unknown parameter matrix. A standard least square algorithm can be applied to estimate the unknown parameter matrix θ' from the sequence of noisy observations $Z_k, k = 1, 2, \ldots, N$.

III. Parameter Estimation via Least Squares

The parameter matrix θ' can be estimated by either a recursive or nonrecursive form. We consider in this article the nonrecursive form. The estimate of θ on the basis of measurement $Z_k, k = 1, 2, \ldots, N$, denoted $\hat{\theta}_N$, is given by

$$\hat{\theta}_N = \left(\sum_{j=1}^{N} x_j x_j' \lambda^{N-j} \right)^{-1} \left(\sum_{j=1}^{N} x_j Z_j' \lambda^{N-j} \right) \qquad (4)$$

where $0 < \lambda \leq 1$ is the exponential data weighting factor. One may refer to Refs. 8 and 11, for example, for an equivalent recursive update of $\hat{\theta}_N$. From $\hat{\theta}_N$, the estimates of A, w_d and ϕ can be obtained.

A. Computational Requirements

The algorithm of Eq. (4) requires an inverse of a symmetric $(n \times n)$ matrix once, requiring order n^2 computations. It may appear that the computation of each $x_j x_j'$ term requires n^2 computations. However, detailed examination shows that only $2n$ computations are required. Thus, the total computations are equal to $6nN + O(n^2)$. In practice, the matrix inverse can be precomputed, thus reducing the data dependent computations to only $2nN + n^2/2$.

B. Fast Implementation of Least Squares Algorithm

The matrix

$$P^{-1} \triangleq \sum_{j=1}^{N} x_j x_j' \lambda^{N-j}$$

in Eq. (4) has a very special structure as can be easily seen by explicit computation of the term $x_j x_j'$ of the summand. Thus,

$$x_j x_j' = \begin{bmatrix} 1 & t_j & t_j^2 & \cdots & t_j^{n-1} \\ t_j & t_j^2 & t_j^3 & \cdots & t_j^{n-1} t_j^n \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ t_j^{n-1} t_j^n & & \cdots & & t_j^{2n-2} \end{bmatrix}$$

Each of the matrices $x_j x_j'$ and P^{-1} is a Hankel matrix. That is, all the elements of each cross-diagonal are the same. The structure of a Hankel matrix is very similar to that of a Toeplitz matrix wherein the elements along the various subdiagonals are equal. The fast algorithm of Ref. 10 for the solution of Toeplitz system of equations can be slightly modified so as to become applicable to the present problem. Thus, the computations in Eq. (4) can be made in order $n(\log_2 n)^2$ computations, resulting in considerable reduction in the requirement for large values of n.

If the matrix inverse is precomputed, then with the algorithmic properties of Ref. 10, the solution for $\hat{\theta}_N$ can be obtained in approximately $6n\log_2 n$ operations. In the implementations above, it is sufficient to store only the first row and column of P or P^{-1}.

C. Baseband Sampling

In the case of baseband sampling, only the measurements $\{y_k\}$ are available and the parameter matrix θ' is of dimension $n \times 1$. In such an implementation, however, there may result an ambiguity of π radians in the phase estimate if the sign of w_d is also unknown.

IV. Estimation Error Analysis

Assuming that the model Eq. (3) is exact (the dimension n of the parameter matrix in Eq. (2) is sufficiently high), then the substitution of Eq. (3) in Eq. (4) yields,

$$\hat{\theta}_N = \left(\sum_{j=1}^{N} x_j x_j' \lambda^{N-j} \right)^{-1} \left\{ \sum_{j=1}^{N} x_j (x_j' \theta + n_j') \lambda^{N-j} \right\} \qquad (5)$$

A simple manipulation of Eq. (5) yields the estimation error $\tilde{\theta}_N \triangleq \theta - \hat{\theta}_N$ as

$$\tilde{\theta}_N = - \left(\sum_{j=1}^{N} x_j x_j' \lambda^{N-j} \right)^{-1} \sum_{j=1}^{N} x_j n_j' \lambda^{N-j} \qquad (6)$$

As the state vector x_j is deterministic, and n_j is a zero mean process, $\tilde{\theta}_N$ has its mean equal to zero. The error covariance matrix of $\tilde{\theta}_N$ can also be evaluated in a straightforward manner. Post multiplying Eq. (6) by the transpose of $\tilde{\theta}_N$, and taking expected values of both sides,

$$E\left[\tilde{\theta}_N \tilde{\theta}'_N\right] = \left(\sum_{j=1}^{N} x_j x'_j \lambda^{N-j}\right)^{-1}$$

$$\times \left\{\sum_{j=1}^{N} \sum_{i=1}^{N} x_j E\left[n'_j n_i\right] x'_i \lambda^{2(N-j)}\right\}$$

$$\times \left(\sum_{i=1}^{N} x_i x'_i \lambda^{N-i}\right)^{-1} \qquad (7)$$

Considering the case of $\lambda = 1$ and recalling that $\{n_j\}$ is a white noise sequence,

$$E\left[\tilde{\theta}_N \tilde{\theta}'_N\right] = \left(\sum_{j=1}^{N} x_j x'_j\right)^{-1} \sigma^2 \qquad (8)$$

$$E\left[\|n_j\|^2\right] = E\left[n_{ij}^2\right] + E\left[n_{qj}^2\right] = \sigma^2$$

A. Frequency Estimation Error

A simple approximate expression can also be obtained for the frequency estimation error when the amplitude A is known and uniform sampling is used. The frequency estimate $\hat{w}_d$ can be obtained as

$$\hat{w}_{d,N} = \left\{\left(\hat{\theta}_N^{21}\right)^2 + \left(\hat{\theta}_N^{22}\right)^2\right\}^{1/2} A^{-1}$$

When the amplitude A is also unknown, it can be replaced by its estimate given by,

$$\hat{A}_N = \left\{\left(\hat{\theta}_N^{11}\right)^2 + \left(\hat{\theta}^{12}\right)^2\right\}^{1/2}$$

In the above expression, $\hat{\theta}_N^{ij}$ denotes the (i,j)th element of the parameter matrix θ_N. The error variance of these elements of interest is given by

$$E\left[\left(\tilde{\theta}_N^{21}\right)^2\right] + E\left[\left(\tilde{\theta}_N^{22}\right)^2\right] \cong K\sigma^2 \left(\sum_{j=1}^{N} t_j^2\right)^{-1}$$

where K approaches a constant with the increase in the numbers of observations N.

For relatively small errors, the frequency estimation error $\tilde{w}_{d,N} = w_d - \hat{w}_{d,N}$ has variance of approximately

$$K \frac{\sigma^2}{2} \left(\sum_{j=1}^{N} t_j^2\right)^{-1}$$

For the case of uniform sampling $t_j = jT_s$, where T_s is the sampling period. Substituting for t_j and letting $T = N T_s$ denote the observation period,

$$E\left[\tilde{w}_N^2\right] = \frac{\sigma^2}{2} K \frac{6}{N(N+1)(2N+1)} \frac{1}{T_s^2} \frac{1}{A^2}$$

In terms of the unsampled system, if the additive noise process has one-sided noise spectral density N_o, then $\sigma^2 = 2N_o/T_s$. Thus,

$$E\left[\tilde{w}_N^2\right] \cong \frac{N_0}{P} \frac{6}{T^3} \frac{K}{4}, \qquad T = NT_s \qquad (9)$$

where $P = A^2/2$ is the received signal power and K has value approximately equal to 4 for low values of n. This is the same mean square error as for Maximum Likelihood estimation (Ref. 1; Ref. 12, Eq. 8.116).

We note here that in the derivation of Eq. (9), the approximation error in Eq. (3) has been ignored. It is difficult to estimate the error due to such finite approximation. However, from a few computer simulations, it appears that for $n > w_d T = (w_d T_s) N$, such error is small.

B. Examples of Application to the DSN Receiver

To keep the dimension n of the parameter matrix small, the following estimation method is proposed. Dividing both sides of Eq. (9) by $w_{d,\max}^2$, and substituting $T = n/w_{d,\max}$, one obtains

$$\frac{E\left[\tilde{w}_d^2\right]}{w_{d,\max}^2} = 6 \frac{N_o}{P} \frac{w_{d,\max}}{n^3}$$

Selecting a value of 1/36 for the left hand side of the above equations allows one to express the maximum frequency uncertainty that can be resolved by the algorithm as a function of n. Thus

$$w_{d,\max} = \frac{n^3}{216} \frac{P}{N_o}$$

The rationale for selecting the value of 1/36 for $E\left[\tilde{w}_d^2\right]/w_{d,\max}^2$ is as follows. Since the additive noise has Gaussian distribution, one may assume that the frequency estimation error has Gaussian distribution with its standard deviation denoted by $\sigma_{\tilde{w}_d}$. The above selection thus ensures that $3\sigma_{\tilde{w}_d} < w_{d,\max}/2$.

Example 1. For reception of Voyager 2 signals at DSS 13, a typical carrier power-to-noise spectral density ratio is 24.4 dB-Hz. Let $n = 8$, and $w_{d,\max} = 652$ rad/s. After an initial

estimation period of $T = n\, w_{d,\max}^{-1}$, the receiver NCO frequency is adjusted by $\widehat{w}_d$. Thus, with an initial adjustment after $T = 12.2$ ms, the frequency offset is reduced to $\widetilde{w}_d$ with $\sigma = 108.6$ rad/s (17.4 Hz). Application of the algorithm for a subsequent period of 24.4 ms reduces the standard deviation to 6 Hz. In this manner, four applications of the algorithm bring down the standard deviation of the frequency offset to less than 0.7 Hz in a total time of 183 ms, from an initial frequency offset of 104 Hz.

Example 2. If the initial uncertainty is only 20 Hz, then with a lower value of n equal to 5, after an estimation period of $T = 40$ ms, $\sigma_{\widetilde{w}_d} = 18.4$ rad/s (2.94 Hz). A frequency correction at the end of this period and an estimation of the residual frequency offset for a period of 160 ms reduces $\sigma_{\widetilde{f}_d}$ to 0.36 Hz ($\widetilde{f}_d = \widetilde{w}_d/2\pi$). Thus with $n = 5$, the frequency uncertainty is reduced to $\sigma_{\widetilde{f}_d} = 0.36$ Hz in a total estimation period of 0.2 s.

In an alternative approach to keep the value of n fixed and small with an increase in the total observation period, instead of resetting the frequency reference (making a correction in the NCO frequency), the time reference is reset to zero. Subsequent observations in Eq. (1) are now with respect to a different phase reference, say $\overline{\phi}$. The application of the least squares algorithm to this set of observations then provides an estimate for $\overline{\phi}$ denoted $\widehat{\overline{\phi}}$. At this stage the observations in the second T s interval are processed to have a phase reference $\widehat{\phi}$ and are then combined with the first set of observations. Equivalently, it is required to post multiply the second sum on the right hand side of Eq. (4), obtained for the second subinterval of T s, by the following matrix

$$\begin{bmatrix} \cos(\Delta\phi) & -\sin(\Delta\phi) \\ \sin(\Delta\phi) & \cos(\Delta\phi) \end{bmatrix}, \quad \Delta\phi = \phi - \widehat{\phi}$$

and add the result to the corresponding sum for the first T s interval. The first sum on the right hand side of Eq. (4) is simply multilied by a factor of 2. This procedure is extended in an appropriate manner to subsequent intervals, so as to obtain a final estimate for w_d, and ϕ based on the complete set of observations.

V. Simulations

Figures 1 through 5 present the frequency estimates obtained by the least squares algorithm. To avoid singularity of the matrix P^{-1}, it was modified by the addition of a diagonal matrix ϵI with $\epsilon = 0.001$. For convenience, the unknown frequency w_d is taken to be 1 rad/s. From the simulations it is apparent that the frequency estimate comes close to the true frequency in a time equal to a fraction of the time period of

the unknown frequency. To keep the computational burden of the simulations to a minimum, the dimension n was restricted to a small value and the observation period was also restricted to a small value.

For frequencies much higher than one, the least squares algorithm, Eq. (4), was slightly modified. Thus, as $w_{d,\max}$ denotes an upper bound on the magnitude of unknown frequency, we define a normalized parameter matrix $\overline{\theta}$ by $\overline{\theta}^{i,j} = \theta^{i,j}/w_{d,\max}^{(j-1)}$; $i = 1, \ldots, n$; $j = 1, 2$. Defining a corresponding state vector $\overline{x}_k$ by

$$\overline{x}_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & w_{d,\max} & \cdots & 0 \\ 0 & 0 & \cdots & w_{d,\max}^{n-1} \end{bmatrix} x_k$$

the measurement Eq. (3) may be rewritten as

$$Z_k = \overline{\theta}' \, \overline{x}_k + n_k \qquad (10)$$

and the least squares algorithm can now be applied to estimate $\overline{\theta}$. The estimates of the elements of θ are then obtained as

$$\widehat{\theta}^{i,j} = w_{d,\max}^{(i-1)} \, \widehat{\overline{\theta}}^{i,j}$$

Such a transformation leaves the previous error analysis invariant. However, for finite dimensional approximation considered here, this makes the algorithm numerically more robust. The simulations for $w_d = 10$ and $w_d = 100$ are precisely the same as in Figs. 1 through 5 with appropriate changes in scaling and are not presented separately.

In Figs. 1 through 5, ψ and ξ represent the first and second row of the parameter matrix respectively. Thus in baseband sampling, only the ψ vector is estimated while in quadrature sampling, the estimates of both ψ and ξ parameter vectors are available. From the figures it is apparent that the dimension n of the state vector $\overline{x}_k$ in the model Eq. (3) is approximately equal to $w_d T$ where T is the sampling interval. Due to over parameterization involved in the problem, there is a considerable amount of flexibility in the estimation of A, w_d, ϕ from the estimates of the elements of θ. Thus whereas in Figs. 1 through 3, the amplitude A is assumed known, Figs. 4 and 5 involve unknown A. A different order of computation can provide an estimate of A when baseband sampling is used. Here, we have reported results only for the frequency estimates; the phase esimates also converge at a fast convergence rate.

VI. Comparison with FFT Techniques

An alternative technique for the fast frequency acquisition is via fast Fourier transform of sampled data. We observe that for the case of infinite observation time, both procedures are optimum and thus are equivalent. However, for finite observation period T, the FFT has the limitation that the frequency estimates are quantized to intervals of $1/T$ Hz. In the finite dimensional approximation of the LS algorithm this is not the case and sufficiently accurate estimates can be obtained by choosing n sufficiently large (finer sampling) even for low values of T.

The price for such an improvement is increased computational requirements which is of order $n\log_2 n$ (though higher than for FFT) if the matrix P is precomputed and is of order $n\log n\ \log n$ if P must be computed on line. With the application of fast algorithms, the storage requirement of P is only $2n$ (not $n^2/2$).

Also, note that the computational requirements here are dominantly decided by $w_d T$ and not by the number of samples as is the case with FFT.

It may also be mentioned that with the FFT algorithm there also exists a finite probability of the occurrence of an outlier (Ref. 1) and this causes a component of the frequency estimation error with a uniform probability density function over the complete frequency range of the FFT algorithm. As against this, the frequency estimation error with LS algorithm has a Gaussian distribution.

VII. Conclusion

This article has presented a fast algorithm based on the least squares parameter estimation technique. In Ref. 9 it is shown that the least squares algorithm exhibits a convergence phase wherein the convergence rate is factorial (the estimation error goes to 0 as $1/k!$ where k is the number of observations) followed by an exponential convergence rate. Our simulations also exhibit the same rapid initial convergence rates. Here of course, the estimation error does not approach zero because of a finite and low dimensional truncation of the model. From another viewpoint the algorithm may be perceived as a time domain dual of the FFT algorithm. Whereas the FFT algorithm transforms the data into frequency domain for the estimation/detection purpose, here the estimation is done directly in the time domain. This latter approach has several advantages. First by choosing $\lambda < 1$, it is possible to track the time varying frequency by recursive update techniques (Ref. 8). Moreover unlike the case of the FFT algorithm, the frequency estimates are not quantized to intervals of $1/T$ Hz, which would be large for small observation interval T. The price for these desirable features is in terms of increased computational requirement which in fast implementation of the algorithm could be of order $n\log n$ or $n\log n\ \log n$ (depending upon the specific implementations), where n is approximately equal to $w_d T$, the product of the frequency uncertainty and the observation period.

References

1. Rife, D. C., Single Tone Parameter Estimation from the Discrete-Time Observations, *IEEE Transactions on Information Theory*, Vol. IT-20, No. 5, pp. 591–598, September 1974.

2. Hurd, W. J., Statman, J. I., and Vilnrotter, V. A., High Dynamic GPS Receiver Validation Demonstration Final Report, *JPL Publication 85-74*, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Calif., 1985.

3. Tufts, D. W., and Kumaresan, R., Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood, *Proceedings of the IEEE*, Vol. 70, No. 9, pp. 975–987, September 1982.

4. Friedlander, B., On the Cramer-Rao Bound for Time Delay and Doppler Estimation, *IEEE Transactions on Information Theory*, Vol. IT-30, No. 3, pp. 575–580, May 1984.

5. Friedland, B., Optimum Steady-State Positions and Velocity Estimation Using Sampled Position Data, *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 9, pp. 906–911, November 1973.

6. Polak, D. R., and Gupta, S. C., Quasi Optimum Digital Phase-Locked Loops, *IEEE Transactions on Communications*, Vol. 21, pp. 75–82, January 1973.

7. Kumar, R., and Hurd, W. J., A Class of Optimum Digital Phase Locked Loops, *TDA Progress Report 42-83*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 63–80, July 1985; also submitted for publication to *IEEE Transactions on AES*.

8. Kumar, R., Almost Sure Convergence of Adaptive Identification and Prediction Algorithms, accepted for publication in *Richerche di Automatica*.

9. Gitlin, R. D., and Magee, F. R., Jr., Self-Orthogonalizing Adaptative Equalization Algorithm, *IEEE Transactions on Communications*, Vol. COM-25, pp. 666–672, July 1977.

10. Kumar, R., A Fast Algorithm for Solving a Toeplitz System of Equations, *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. ASSP-33, No. 1, pp. 254–267, February 1985.

11. Kumar, R., Simultaneous Adaptive Control and Identification via Weighted Least Squares Algorithm, *IEEE Transactions on Automatic Control*, Vol. 29, pp. 259–263, March 1984.

12. Wozencraft, J. M., and Jacobs, I. M., *Principles of Communication Engineering*, New York: Wiley, 1965.
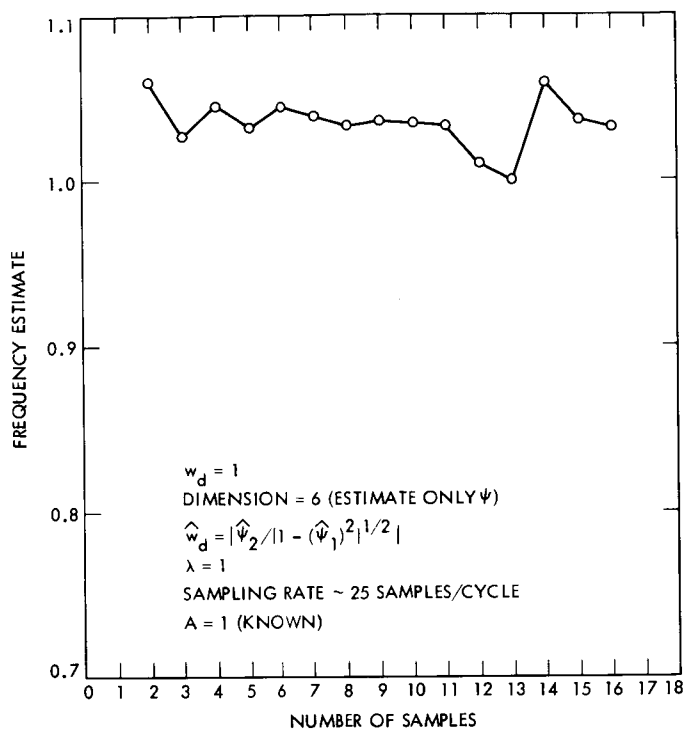
Fig. 1. Least squares algorithm: noise free case and baseband sampling, $n = 6$
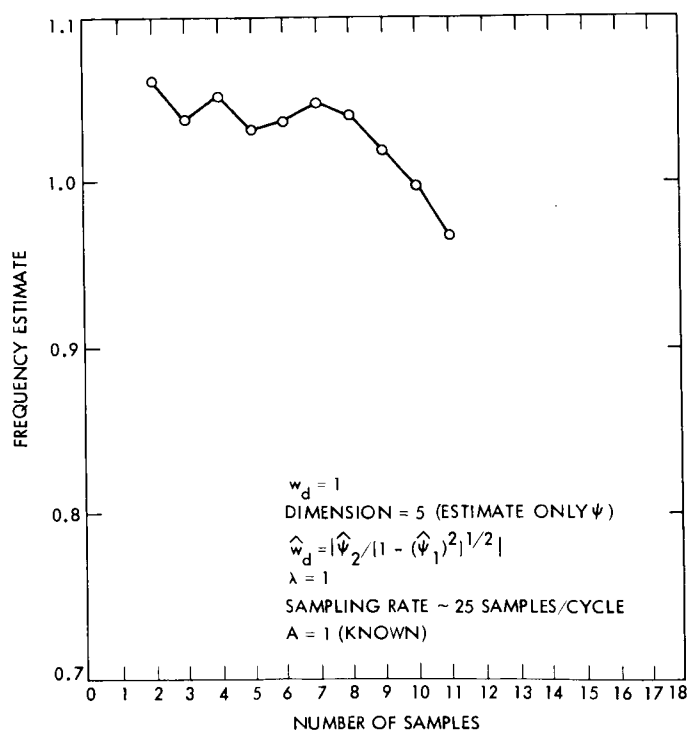
Fig. 2. Least squares algorithm: noise free case and baseband sampling, $n = 5$
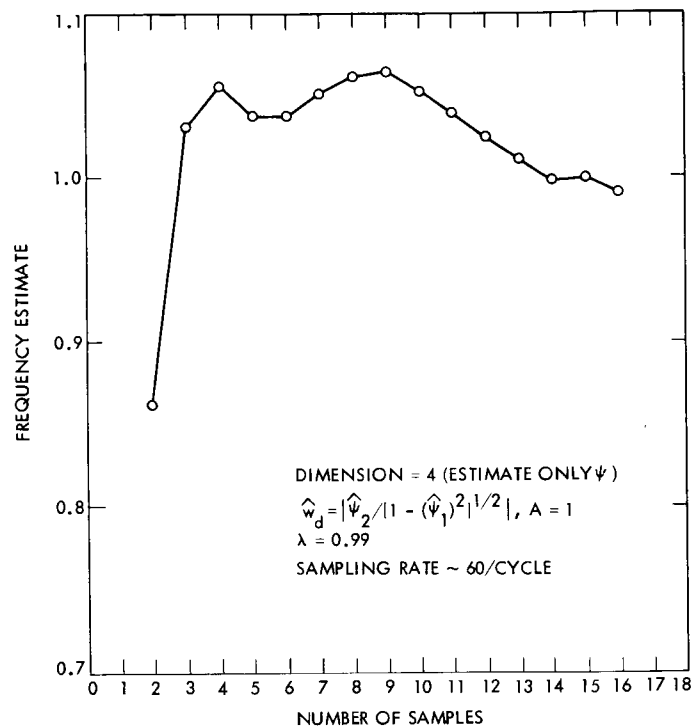
Fig. 3. Least squares algorithm: noise free case and baseband sampling, $n = 4$
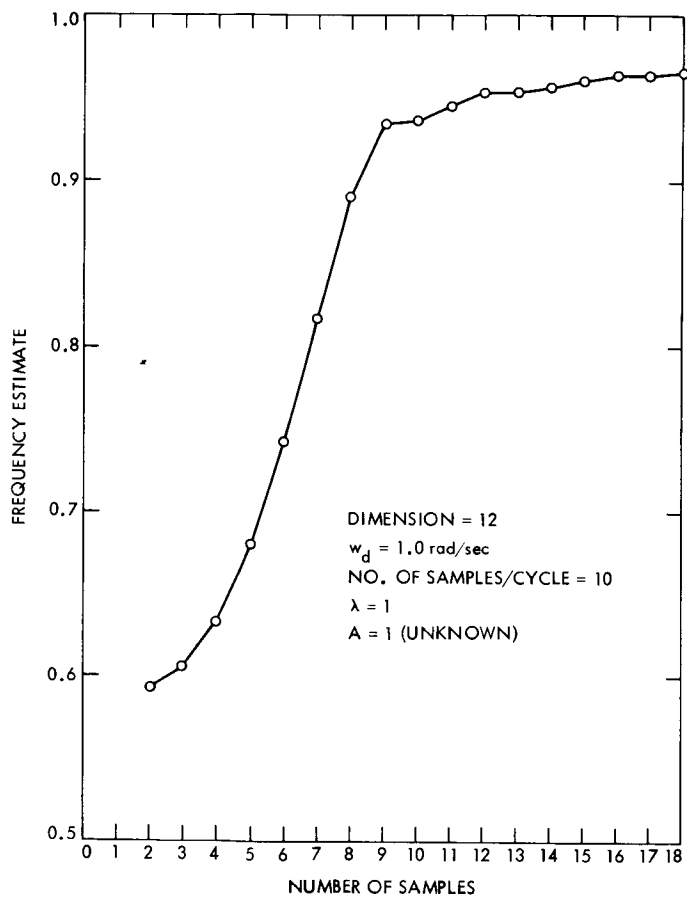
Fig. 4. Least squares algorithm: noise free case and quadrature sampling

Graph text:
FREQUENCY ESTIMATE (y-axis), 0.5 to 1.0
NUMBER OF SAMPLES (x-axis), 0 to 18

DIMENSION = 12
w_d = 1.0 rad/sec
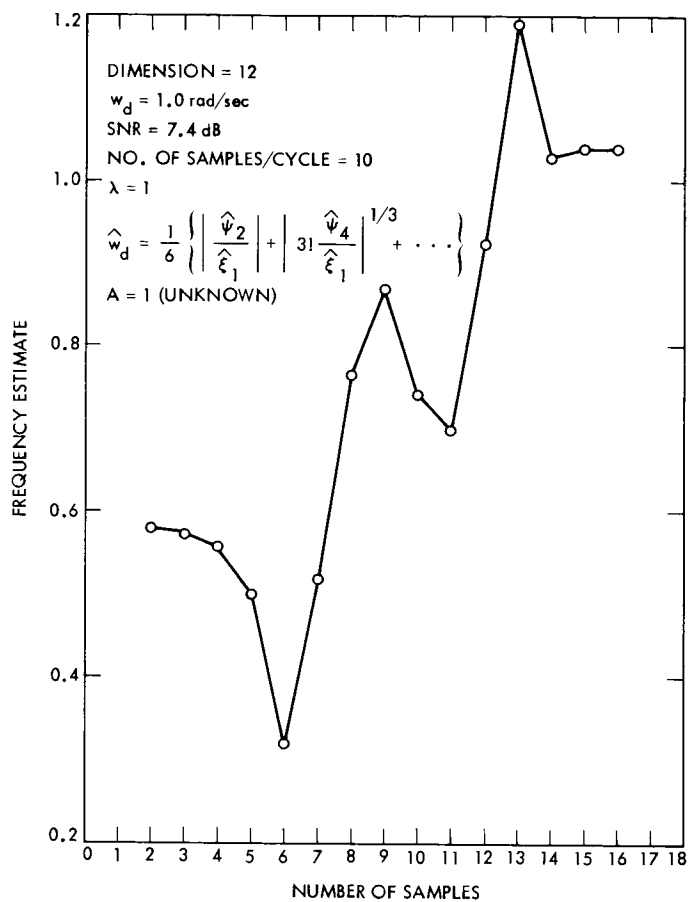NO. OF SAMPLES/CYCLE = 10
λ = 1
A = 1 (UNKNOWN)

Fig. 5. Least squares algorithm and quadrature sampling

Graph text:
FREQUENCY ESTIMATE (y-axis), 0.2 to 1.2
NUMBER OF SAMPLES (x-axis), 0 to 18

DIMENSION = 12
w_d = 1.0 rad/sec
SNR = 7.4 dB
NO. OF SAMPLES/CYCLE = 10
λ = 1

$$\widehat{w}_d = \frac{1}{6}\left\{\left|\frac{\widehat{\psi}_2}{\widehat{\xi}_1}\right| + \left|3!\frac{\widehat{\psi}_4}{\widehat{\xi}_1}\right|^{1/3} + \cdots\right\}$$

A = 1 (UNKNOWN)